Predictor Pro: A Unified Machine Learning Application for Stock Market and Weather Prediction

Pole Vennela [1], M. Mary Sujatha[2], G. Sreedhar[3]
[1]*Student, M. Sc Computer Science, Department of Computer Science, National Sanskrit University, Tirupati – 517507, Andhra Pradesh*
[2]*Assistant Professor, Department of Computer Science, National Sanskrit University, Tirupati – 517507, Andhra Pradesh.*
[3]*Professor, Department of Computer Science, National Sanskrit University, Tirupati – 517507, Andhra Pradesh*

Abstract: Predictor Pro introduces a desktop application. It demonstrates prediction of stock market movements and weather conditions using machine learning algorithms. Predicting stock market trends is difficult because of market volatility and the complexity of financial time series data. Traditional statistical methods often struggle to capture the nonlinear relationships that affect price changes. Weather forecasting is important task in agriculture, transportation, and disaster management. Predictor Pro is a desktop application that applies six machine learning algorithms namely Logistic Regression, K-Nearest Neighbour, Decision Trees, Random Forest, Support Vector Machines, and Gradient Boosting. It predicts stock price movements and forecasts the weather. The system analyses historical price data, including Open, High, Low, Close, and Volume, to predict if the next day's closing price will rise (UP) or fall (DOWN). For weather forecasting, it processes historical data like Date, Precipitation, Maximum temperature, Minimum temperature, Wind, and Weather to forecast the following day's weather conditions and provide recommendations. Experimental results indicate that the overall accuracy varies across different algorithms, with accuracy levels ranging from 55% to 65%. Support Vector Machines and Decision Trees perform better than the other methods. The application features an easy-to-use GUI, collects real-time data through the Yahoo Finance API, offers interactive visualizations, and provides live predictions. Although it is not suitable for actual trading decisions, Predictor Pro is a useful educational tool for learning about machine learning in financial forecasting.

Keywords: Machine learning, Stock Market, Weather Forecasting, Support Vector Machine (SVM), Decision Tree, Random Forest, Gradient Boosting, K-Nearest Neighbour (KNN), Logistic Regression.

1.INTRODUCTION

Prediction techniques are the tools and methods employed in the estimation of future outcomes based on historical and present data. These techniques are very important in various fields, such as finance, weather forecasting, healthcare, business intelligence, and engineering as shown in the Figure1. The fundamental goal behind prediction is to analyse historical data, establish relationships between variables, and make accurate predictions about future occurrences [1].

In today's data-driven world, accurate prediction systems play a crucial role in decision-making across multiple sectors. Financial markets and weather systems are two highly dynamic and complex domains that significantly impact economies, businesses, and daily life [2]. Predictor Pro presents an integrated machine learning-based system that combines Stock Market Prediction and Weather Forecasting into a unified platform.

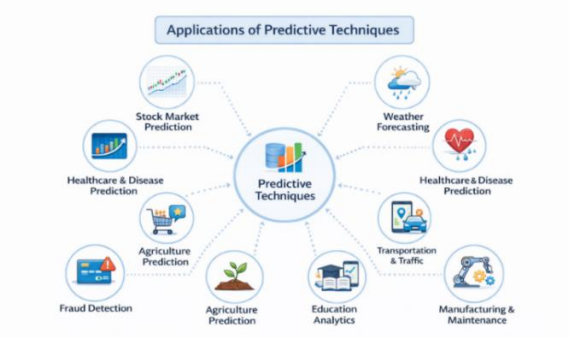

Figure 1: Applications of Predictive Techniques

Traditionally, prediction is done using statistical models such as Linear Regression, Moving Average,

and Time Series models such as ARIMA. These models are very effective in cases where the relationship between variables is simple and linear [3]. Nevertheless, real-world applications such as stock exchanges and weather conditions are highly complex, nonlinear, and involve multiple factors, which are dynamically interacting with each other. These statistical models cannot handle such complex scenarios.

With the emergence of Artificial Intelligence, Machine Learning and Deep Learning models have become very popular in predictive analysis. Machine learning models can learn from data and make predictions without being explicitly programmed. Decision Trees, Random Forest, Support Vector Machines (SVM), and k-Nearest Neighbours (k-NN) are some examples of machine learning algorithms employed in prediction models [4][5].

Illustration of Machine Learning algorithms in prediction of stock market and weather forecasting is possible by training algorithms with historical data and testing today and tomorrow values [6]. The Machine learning algorithms that we have used are discussed in the following grounds.

Linear Regression: It is one of the most commonly used supervised machine learning algorithms for prediction problems. It establishes a relationship between independent variables and a dependent variable by fitting a linear equation to the observed data. In stock market prediction and weather forecasting, Linear Regression helps estimate future values based on historical trends.

Decision Tree: It is a supervised machine learning algorithm used for both classification and regression tasks. It works by dividing the dataset into smaller subsets based on feature values. Each internal node represents a decision based on a feature, and each branch represents the outcome of that decision. In predictive analytics, Decision Trees help identify patterns in stock price trends and weather variables. The algorithm is easy to interpret and can handle complex datasets with multiple features.

Random Forest: It is an ensemble machine learning algorithm that combines multiple Decision Trees to improve prediction accuracy. Each tree is trained on a random subset of the dataset, and the final prediction is obtained by averaging the results of all trees. Random Forest helps reduce overfitting and improves prediction reliability. Predict Pro used it to analyse patterns in stock market data and weather parameters.

Support Vector Machine: It is a powerful machine learning algorithm used for both classification and regression problems. SVM is effective in handling high-dimensional datasets and can model complex relationships between variables.

Logistic Regression: It is a supervised machine learning algorithm commonly used for classification problems. It is used to predict the probability of a categorical outcome based on one or more independent variables. In predictive systems, Logistic Regression can be used to classify outcomes such as predicting whether a stock price will increase or decrease, or determining whether certain weather conditions such as rainfall may occur.

Gradient Boosting: It is used in Predictor Pro to improve prediction accuracy for both stock market trends and weather forecasting. The algorithm analyses historical datasets and learns complex relationships between different features. By sequentially correcting prediction errors, the model produces more accurate forecasts compared to single machine learning models.

K-Nearest Neighbour: KNN is used in Predictor Pro to predict future stock prices and weather conditions by comparing new data points with historical data. The algorithm identifies patterns from similar past observations and generates predictions based on those similarities.

II. LITERATURE REVIEW

Predictive techniques have gained significant attention in recent years due to the rapid growth of data availability and advancements in machine learning techniques. Stock market prediction primarily relied on statistical techniques such as Linear Regression, Moving Averages, and ARIMA models [7]. These methods focused on analysing historical price trends and identifying linear relationships among financial variables.

According to *Siddesh Kokare et al* [8], various Machine Learning models such as Long Short Term Memory (LSTM), Convolution Neural Networks (CNN) and Convolution Neural Networks Long Short Term Memory (CNN-LSTM) are using conventional strategic measures such as MAE (Mean, Absolute,

Error). The measured low values indicates that the models are effective in predicting stock prices. The developed model uses stock datasets parameters such as close price, open price, high price, low price, previous closed price, turnover and volume. These parameters are used as the input to the models for training and testing stock price prediction. The results were observed that Convolution Neural Networks - Long Short Term Memory (CNN-LSTM) has performed better than other two models.

Jagruti Hota et al [9] demonstrated four algorithms such as Decision Tree, Support Vector Machine (SVM), Random Forest and Artificial Neural Networks (ANN) were trained with available stock data of American Airlines for the last 5 years. The data was split into 70% for training and 30% for testing. Results showcased that Random Forest performed better than other three models as it has lowest Mean Absolute Percentage Error (MAPE) value of 0.36%. Thus, which can be used in the real-time implementation.

The research done by *Ismaila Oshodi* [10] on machine learning based weather forecasting model where he trained four algorithms on the dataset which is obtained from Kaggle for the city of Seattle for the period 2012 to 2015.The weather parameters that are used to evaluate the model are drizzle, fog, rain, snow, and sun. Algorithms such as Random Forest, Decision Tree, Gaussian Naïve Bayes, and Gradient Boosting Classifier are used for model evaluation where Gaussian Naïve Bayes algorithm gave the best predictive accuracy of 84.153%.

Munmun Biswas et al [11] used combination of Navie Bayes and Chii Square algorithms to predict weather conditions, which is a web application based system with effective graphical User Interface where the user login to the system with User ID and Password. This system will take this parameter and predict weather after analysing the input information with the information in database. Given framework arranges the information into various classifications. This methodology can decide the nonlinear relationship that exists between the historical data (temperature, wind speed, humidity) provided to the system during the training phase and on that premise, make a prediction of what the weather would be in future.

By all these advancements in prediction using Machine learning Algorithms *McCarthy. P* proposed an unified model [12]. in prediction of Stock Market and Weather Forecasting where the system uses historical dataset that is loaded to make predictions. Given system improves the accuracy of models based on algorithms performance on the dataset to make Today and Tomorrow Predictions.

III. METHODOLOGY

The proposed system follows a machine learning-based predictive framework to forecast stock prices and weather conditions. The methodology consists of several stages including data collection, data preprocessing, model training, model evaluation, and prediction generation. Historical datasets from financial markets and meteorological sources are analysed using machine learning algorithms to generate accurate predictions. The workflow of the system is shown in the Figure 2.

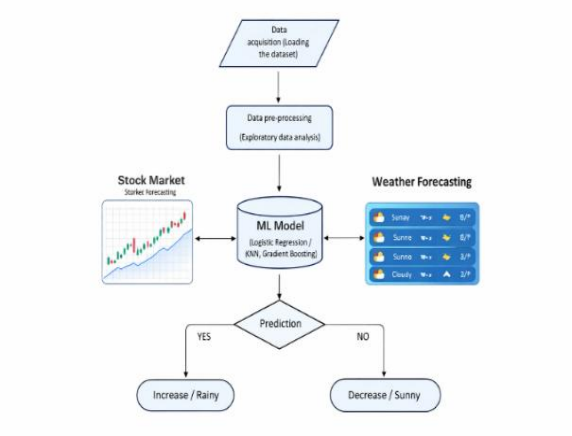

Figure 2: Workflow of combined system for stock and weather forecasting

Step 1: Collecting historical and real-time datasets. Stock Market dataset - S&P 500 (^GSPC) collected publicly from Kaggle which includes historical stock prices including open, high, low, close prices and trading volume from 1927-12-30 to 2026-03-03. Weather forecasting dataset is collected publicly from Kaggle which includes historical weather data including Date, Precipitation, Maximum Temperature, Minimum Temperature, Wind and Weather from 2012-01-01 to 2015-12-31.

Step 2: Data Pre-processing is a next step in the prediction process where the collected datasets may contain missing values, inconsistent records, or noise that can negatively affect the model performance.

Therefore, several pre-processing techniques [13][14] are applied for transformation of raw data into a structured format suitable for machine learning algorithms.

Step 3: Selection of algorithms

Step 4: Model training is another step where the selected algorithms learn patterns and relationships from the historical dataset. During this phase, the model analyses input data and adjusts its internal parameters to make accurate predictions [15]. Before training the model, the dataset is divided into two main parts, typically in 80:20 ratio, where 80% of the data is used for training the model. 20% of the data is used for testing the model to evaluate the model's performance.
During the training process, the algorithms adjust their internal parameters to minimize the prediction error between the predicted values and the actual values. In stock market prediction, the model learns relationships between historical stock prices, technical indicators, and market trends. In weather forecasting, the model learns patterns from meteorological variables such as temperature, humidity, wind speed, and atmospheric pressure.

Step 5: After training the model, validation is performed to ensure that the model has learned meaningful patterns rather than memorizing the training data. The trained model is tested using the testing dataset to evaluate how well it performs on unseen data [16]. Predictor Pro trained multiple machine learning algorithms using historical stock market and weather datasets to develop a robust predictive system capable of generating accurate forecasts.

Step 6: Model evaluation is an important step in the machine learning workflow used to measure the performance and accuracy of the trained model. After the model has been trained using the training dataset, it must be tested using a separate testing dataset to determine how well it can make predictions on new and unseen data. This helps ensure that the model has learned meaningful patterns rather than simply memorizing the training data [17].

Step 7: Once the machine learning model has been trained and evaluated, it is used to generate predictions based on new input data. The prediction process involves applying the trained model to unseen data in order to forecast future outcomes [18].

In the proposed system, the prediction process works as follows:
i) Input Data: New data related to stock market parameters or weather conditions is provided to the system.
ii) Data Pre-processing: The input data undergoes pre-processing steps such as normalization and feature extraction to ensure it is compatible with the trained model.
iii) Feature Extraction: Relevant features are selected from the input data, including stock indicators or weather variables.
iv) Model Prediction: The trained machine learning model processes the input data and generates predictions for future stock prices or weather conditions.
v) Output Generation: The predicted results are displayed to the user in the form of numerical values and graphical visualizations.

IV.. EXPERIMENTAL RESULTS

For stock market prediction, historical stock price data including opening price, closing price, highest price, lowest price, and trading volume were used as input features. The models analysed the trends and patterns present in the dataset to predict future stock price movements.
The results indicates that machine learning algorithms are capable of capturing relationships between market variables and generating reasonable predictions. Among the tested algorithms, Support Vector Machine (SVM) showed better performance in handling nonlinear patterns in the dataset and produced more accurate predictions compared to simpler models. Table 1 & Figure 3 shows the accuracy of the models after training the dataset.

Model	Accuracy (%)
Logistic Regression	54.65 %
K Nearest Neighbour	51.31 %
Decision Tree	52.5 %

Random forest	46.75 %
Support Vector Machine	54.75 %
Gradient Boosting	45.96 %

Table 1: Accuracy of Models for Stock Market Prediction

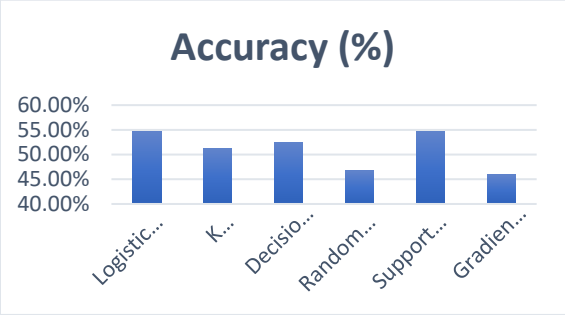

Figure 3: Accuracy of Stock Market Prediction

The predicted stock prices were compared with the actual stock prices using graphical visualization techniques such as line graphs shown in the Figure 4. The graph showed that the predicted values closely follow the actual market trends, indicating the effectiveness of the model.

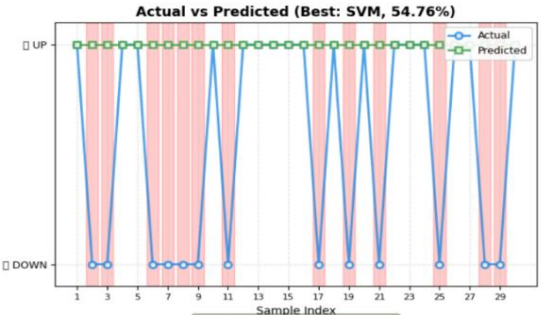

Figure 4: Actual vs Predicted Stock Prices

For weather forecasting, the system used meteorological data such as Date, Precipitation, Maximum Temperature, Minimum Temperature, Wind and Weather. The models analysed historical weather patterns to forecast future weather conditions. The experimental results demonstrate that machine learning models can effectively identify patterns in weather datasets and provide reliable predictions for short-term forecasting. Among the tested algorithms, Support Vector Machine (SVM) showed better performance to produce more accurate predictions compared with other models. Table 2 & Figure 5 shows the accuracy of the models after training the datasets.

Model	Accuracy (%)
Logistic Regression	80.2 %
K Nearest Neighbour	77.82 %
Decision Tree	69.97 %
Random forest	79.18 %
Support Vector Machine	81.23 %
Gradient Boosting	78.16 %

Table 2: Accuracy of Models for Weather Forecasting Prediction

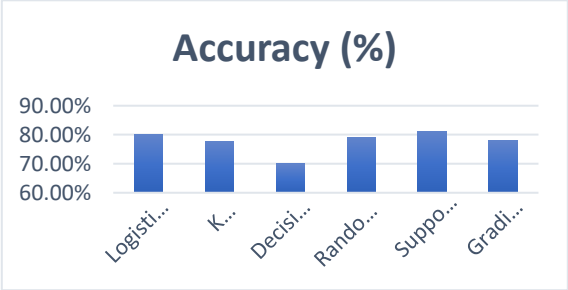

Figure 5: Accuracy of Weather Forecasting Prediction

The predicted weather parameters were compared with actual values to evaluate the accuracy of the models as shown in Figure 6. The visualization of weather predictions through graphs and charts helped in understanding variations in temperature and other weather variables over time.

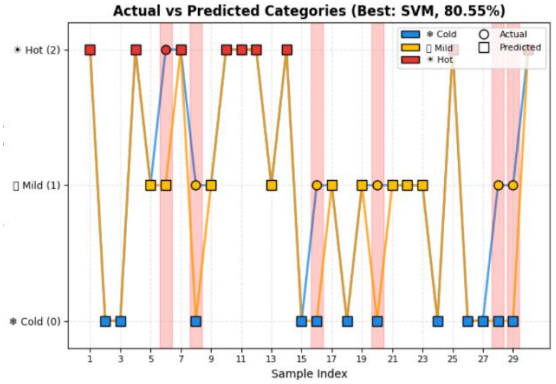

Figure 6: Actual vs Predicted Weather Forecasting

V.CONCLUSION AND FUTURE WORK

The developed system successfully predicts stock market trends and weather conditions with reasonable accuracy, proving the effectiveness of the proposed approach. The models use stock dataset parameters such as opening price, closing price, high price, low

price, volume and weather dataset parameters such as Date, Precipitation, Maximum temperature, Minimum temperature, Wind and Weather. These parameters are used as the input to the model for training and testing for stock prices and weather forecasting predictions. The experimental results show that in most of the cases, the Support Vector Machine (SVM) performs very much better than other models with highest accuracy of 50 - 80 %. The model provides results based on historical data which predicts today and tomorrow predictions. Future work includes adding of another mode such as Election result prediction, Cricket match prediction, Human Disease prediction and Sales prediction etc., to the proposed application so that it can also provide the results in the same application. This work may also increase the accuracy of the models to predict stock prices and weather conditions by considering historical data.

REFERENCES

[1] Vaibhav Kumar, M.L.Garg, Predictive Analytics: A Review of Trends and Techniques (2018), Volume 182 -No.1.

[2] Polamuri Subba Rao, K. Srinivas, and A. Krishna Mohan, A Survey on Stock Market Prediction Using Machine Learning Techniques (2020), https://doi.org/10.1007/978-981-15-1420-3_101

[3] Hegazy O, Soliman OS, Salam MA (2013) A machine learning model for stock market prediction. Int J Comput Sci Telecommun (IJCST) 4(12)

[4] Ghosh et al., "Weather Data Mining using Artificial Neural Network," 2011 IEEE Recent Advances in Intelligent Computational Systems, Trivandrum, 2011, pp. 192-195.

[5] Introduction to Data Mining and Knowledge Discovery, Third Edition, Two Crowds Corporation,http://www.twocrows.com/introdm. pdf, accessed on 12 April 2009.

[6] Hewage, P., Trovati, M., Pereira, E. and Behera, A. (2021). Deep Learning-Based Effective Fine-grained Weather Forecasting Model. Pattern Analysis and Applications, 24(1), 343-366.

[7] Kaggle. (2022). Weather Prediction [online]. Kaggle.com. Available from: https://www.kaggle.com/datasets/ananthr1/weather-prediction/code [Accessed May 10, 2022].

[8] Siddesh Kokare, Anvit Kamble, Shubham Kurade and Deepali Patil, Study And Analysisof Stock Market Prediction Techniques (2022), https://doi.org/10.1051/itmconf/20224403033

[9] Jagruti Hota, Sujata Chakravarty, Bijay K. Paikaray and Harshvardhan Bhoyar, Stock Market Prediction Using Machine Learning Techniques (2022), CEUR-WS, Vol-3283.

[10] Ismaila Oshodi, Machine Learning-Based Algorithms for Weather Forecasting (2022), International Journal of Artificial Intelligence and Machine Learning, 2(2), 12-20. doi: 10.51483/IJAIML.2.2.2022. 12-20.

[11] Munmun Biswas, Tanni Dhoom, and Sayantanu Barua, Weather Forecast Prediction: An Integrated Approach for Analysing and Measuring Weather Data (2018), Volume 182 – No. 34.

[12] McCarthy, P. (2007). Defining the Impact of Weather. In Proc. 22nd Conference on Weather Analysis and Forecasting/ 18th Conference on Numerical Weather Prediction, American Meteorological Society, Utah.

[13] Deepak, Raut Sushrut, Shinde Isha Uday, and D. Malathi. "Machine learning approach in stock market prediction." International Journal of Pure and Applied Mathematics 115.8 (2017): 71 77.

[14] Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5-32.

[15] Vanukuru, Kranthi. (2018). Stock Market Prediction Using Machine Learning. 10.13140/RG.2.2.12300.7748.

[16] Aparna Nayak, M. M. Manohara Pai and Radhika M. Pai, “Prediction Models for Indian Stock Market”, (IMCIP) (2016).

[17] Samarawickrama AJP, Fernando TGI (2017) A recurrent neural network approach in predicting daily stock prices, An application to the Sri Lankan stock market. IEEE

[18] Mehtabhorn Obthonga, Nongnuch Tantisantiwong, Watthanasak Jeamwatthanacha c and Gary Willsd, “A Survey on Machine Learning for Stock Price Prediction: Algorithms and Techniques”, (2020).

[19] Wenjie Lu, 1, 2 Jiazheng Li, 3 Yifan Li, 3 Aijun Sun, and Jingyang Wang, “A CNN-LSTM-Based Model to Forecast Stock Prices”, (2020).